Secondary Students' Reasoning on Pedigree Problems

Justin Timm,* Katharina Wools, and Philipp Schmiemann
Biology Education Research and Learning Lab, University of Duisburg-Essen, 45141 Essen, Germany

ABSTRACT

Pedigree problems are typical genetics tasks in schools. They are well suited to help students learn scientific reasoning, representing realistic genetic problems. However, pedigree problems also pose complex requirements, especially for secondary students. They require a suitable solution strategy and technical knowledge. In this study, we examined the approaches used by *N* = 89 secondary school students when solving two different pedigree problems. In our qualitative analysis of student responses, we examined how two groups of secondary students with varying degrees of experience in genetics constructed arguments to support their decisions. To do so, we categorized *I* = 516 propositions from students' responses using theory- and data-driven codes. Comparison between groups revealed that "advanced genetics" students (*n* = 44) formulated more arguments, referred more frequently to specific family constellations, and considered superficial pedigree features less often. Conversely, "beginning genetics" students did not use a conclusive approach of step-by-step falsification but argued for the mode of inheritance they believed was correct. Advanced genetics students, in contrast to beginners, to some extent used a falsification strategy. Finally, we demonstrate which family members students used in their decisions and discuss a variety of typical but unreliable arguments.

INTRODUCTION

Pedigrees or family trees are a highly standardized graphical representation of family medical histories (Bennett *et al.*, 1993, 1995, 2008; Bennett, 2010). They are used in genetic counseling and research to determine genetic influence on a trait, to identify patterns of inheritance, to calculate the disease risk for an individual, and for multiple other reasons (Bennett, 2010, presents an extensive list). Moreover, "human pedigree analysis is a typical example of genetics problem solving" (Corbett *et al.*, 2010, p. 221). In this paper, we analyze secondary students' argumentation and strategic approaches when solving pedigree problems, as these are well suited for learning scientific reasoning. We used a qualitative approach and compared two groups of secondary students with divergent levels of genetics education.

Pedigree Problems

Pedigree problems are frequently used in genetics education. A typical pedigree problem consists of a representation of a human genetic pedigree with a task aim set to identify the mode of inheritance that is (most likely) represented. Typically, four modes of inheritance are differentiated: autosomal dominant inheritance, autosomal recessive inheritance, X-linked dominant inheritance, and X-linked recessive inheritance.

In autosomal dominant inheritance, the presence of a single dominant allele results in the expression of the trait for which the allele codes. Each person with the trait has at least one affected parent. In autosomal recessive inheritance, the trait is expressed only if a person has inherited one allele from each parent. Heterozygous individuals do not develop the trait. They can, however, still inherit the affected allele and are, therefore, called carriers. An X-linked dominant trait is encoded by an allele located on

Jennifer Knight, *Monitoring Editor*
Submitted Jan 20, 2021; Revised Dec 17, 2021; Accepted Jan 3, 2022
CBE Life Sci Educ March 1, 2022 21:ar14
DOI:10.1187/cbe.21-01-0009
*Address correspondence to: Justin Timm (justin.timm@uni-due.de).

the X chromosome. Males have only one X chromosome. As a result, all their daughters but none of their sons will inherit the trait. For females, however, both sons and daughters can inherit the trait. In X-linked recessive inheritance, only the absence of an unaffected allele results in the expression of the trait. Males are more likely to be trait carriers, because they cannot compensate for an inherited allele.

A complete solution to a pedigree problem requires the identification of the underlying mode of inheritance; it consists of a decision in favor of one pattern and, in the best case, evidence for the exclusion of all alternative patterns. Therefore, an exhaustive solution requires testing of multiple hypotheses (Hackling and Lawrence, 1988; Smith and Good, 1984; Hackling, 1994). In a systematic approach, all modes of inheritance are tested to determine whether they can explain the trait's distribution. Basically, this process constitutes a search for inconsistencies in which family constellations are identified as cues that would not be possible assuming one of the modes of inheritance in question; in this context, we define the genetic family constellation as the elementary informative unit consisting of father, mother, and one child. An inconsistency such as an affected child with unaffected parents, for example, proves that the trait is not dominant. Smith (1988) calls the combination of such an informative family constellation and the conclusion to be drawn from it a production rule. These production rules consist of condition–action pairs for which identifying a condition—for example, the occurrence of unaffected parents with an affected child—leads to a specific action, in this case, the conclusion that the trait cannot be dominant (Smith, 1988). In the case of a pedigree problem with a clear solution, all but one mode of inheritance can be excluded by using a combination of several of these rules or by systematically assigning genotypes to all individuals and looking for contradictions. However, pedigree problems seem to be challenging for students (Hackling and Lawrence, 1988; Smith, 1988; Hackling, 1994; Knippels *et al.*, 2005). Various difficulties have been described. In the study by Smith (1988), unsuccessful participants tended to consider the first plausible option as an acceptable answer and used an incomplete falsification strategy during pedigree analysis. They rarely tested hypotheses using genotypes and often used imprecise or incorrect logic. They tended to identify a larger number of noncritical cues; they then based decisions on these cues and tended to be unable to make use of production rules. In short, they failed because their rules were incomplete or incorrect, inappropriately applied, or did not lead to any conclusions. In addition, they based decisions on genetic ratios, which is usually not reasonable, as pedigrees are normally too small to comply with statistical distributions. Hackling and Lawrence (1988) demonstrated that university genetics professors' solutions to pedigree problems were more complete than those of both novice and competent human biology university students. The experts (professors) identified a larger number of critical cues and varied their hypothesis-testing strategy according to the prevalence of the trait. They were likely to use genotypes to test hypotheses for a common trait, but relied more on critical cues for rare traits. In contrast, in the case of students, this difference was not noticeable. In another study, Hackling (1994) found that, in a group of novice students, almost no one succeeded in correctly interpreting cues with respect to X-linked inheritance. In addition, a significant proportion of this group was unable to assign genotypes to a pedigree that showed the occurrence of an X-chromosomal dominant trait. In a study by Corbett *et al.* (2010), if-then statements (which they call "cognitive statements") to determine carrier status in case of an X-linked recessive inheritance caused the greatest difficulty for university students.

Pedigree Problems as One Type of Genetic Problem

Pedigree problems are just one type of genetic problem that requires problem-solving skills. They are classified as rather difficult problems, because they require a fairly challenging "effect-to-cause" reasoning between generations, whereas the simplest problem type, in contrast, only requires "cause-to-effect" reasoning within generations (Stewart, 1988; Hickey *et al.*, 2000; Tsui and Treagust, 2010). Pedigree problems belong to the area of Mendelian genetics, which is based on the principles of meiosis (Stewart *et al.*, 2005). Collins and Stewart (1989) identified four typical classes of problems within the area of Mendelian genetics, which they grouped into inheritance pattern problems and modifier problems: Inheritance patterns describe the relationship between genotype and phenotype. Typical patterns are simple dominance, codominance, and multiple alleles (which is a combination of the previous two; Collins and Stewart, 1989). Modifiers affect the inheritance pattern by altering or constraining the way alleles are transmitted. They can be present or absent but are bound to an inheritance pattern; they cannot exist without it. Possible modifiers include X-linkage, genetic linkage (when genes for two traits are closely to each other located on a chromosome), and lethal alleles (Collins and Stewart, 1989). According to Collins and Stewart (1989), the inheritance patterns of simple dominance, codominance, and multiple alleles and X-linkage as a modifier are all typically considered in introductory courses.

In pedigree problems, however, simple dominance and X-linkage are predominantly considered. While in other tasks the modifier "X-linkage" may be present or absent, this does not apply to pedigree problems: Once introduced, all four modes of inheritance (autosomal dominant inheritance, autosomal recessive inheritance, X-chromosomal dominant inheritance, and X-chromosomal recessive inheritance) are almost always considered. As a result, a complete solution should typically take all four modes of inheritance into account. In any case, pedigree problems are well suited for inquiry-based genetics education (Stewart *et al.*, 2005).

Argumentation and Content Knowledge

Tasks involving pedigrees require problem-solving skills, because students need to use content knowledge, formulate hypotheses, and apply one or multiple solution strategies (Aznar and Orcajo, 2005). They need to generate conclusive arguments to prove their problem solutions, and argument quality is presumed to be related to content knowledge (Sadler and Zeidler, 2005; Sadler and Fowler, 2006; McNeill *et al.*, 2006; Aufschnaiter *et al.*, 2008). But what makes a conclusive argument? A simple argument consists of a claim supported by data as evidence (Toulmin, 2003). However, more complex arguments can have a number of other components. Following the structure of Toulmin (2003), arguments can also include warrants and backings, qualifiers, and rebuttals. Warrants are generally applicable rules and laws that may be included to

legitimize the step from data to claim. Backings can be cited to legitimize the applicability of the warrants. As a simplification of the Toulmin model, warrants and backings can be summarized under the term "reasoning" (McNeill *et al.*, 2006; Berland and McNeill, 2010); that is, reasoning encompasses justifying how the evidence chosen leads to the claim made. The model of Toulmin (2003) also contains qualifiers and rebuttals: Qualifiers modify the strength of a claim by taking into account rebuttal(s) in the form of exceptions and conditions under which the warrant does not apply. Transferred to our context, a good and comprehensive argument against a particular inheritance might look like this: The trait at hand cannot be inherited in an autosomal dominant manner (claim), because a couple without the trait has an affected child (evidence). This rules out autosomal dominant inheritance, because at least one parent would also have to be an allele carrier for the child to inherit a dominant allele. However, this parent would also have to be affected, because this dominant allele would have led to the expression of the corresponding phenotype (warrant or reasoning). As a result, to formulate a complete explanation to a pedigree problem, students must formulate conclusive arguments against three of the four relevant modes of inheritance based on content knowledge and evidence available in the pedigree.

Regardless of content knowledge, students struggle to formulate coherent arguments. They seem to have problems citing sufficient evidence (Sandoval and Millwood, 2005) and justifying why evidence leads to specific claims (McNeill *et al.*, 2006). According to three different learning progressions (Songer *et al.*, 2009; Berland and McNeill, 2010; Lee *et al.*, 2014), students' scientific argumentations and explanations can be categorized according to their complexity: Basal scientific argumentations and explanations are based on unsubstantiated claims, while more elaborated explanations and argumentations are characterized by claims that are justified by evidence. Arguments that entail warrants and backings (reasoning) on how evidence and assertion are related indicate an even more complete or complex level of scientific argumentation or explanation.

According to the model of Shea *et al.* (2015), the interplay of content knowledge, argumentation skills, and situational features accounts for genetics literacy. Therefore, situational features may also influence genetic reasoning and thus problem solving (Freidenreich *et al.*, 2011; Shea *et al.*, 2015; Schmiemann *et al.*, 2017).

Purpose

To clearly determine the mode of inheritance in the case of a pedigree problem, all alternative modes of inheritance must be excluded with justification. The literature suggests that even university students often struggle with pedigree problems and fail to form conclusive and complete solutions. Nevertheless, students are already expected to solve pedigree problems in secondary school. With this study, we wanted to explore secondary school students' approaches to solving pedigree problems and how they constructed arguments to support their decisions. Our research questions were: What arguments do secondary school students generate when analyzing human genetic pedigrees and what evidence are the students' arguments based on? Do their arguments differ depending on their level of genetics education?

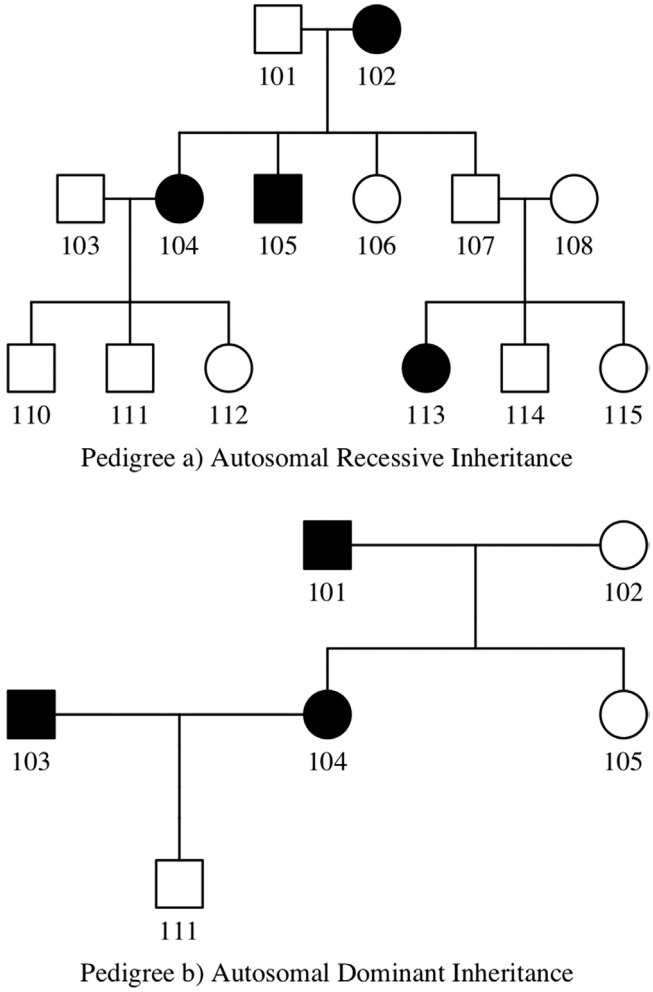

FIGURE 1. Pedigrees used in the paper-and-pencil tests.

METHODOLOGY

Instrument and Data Collection

In pursuit of our aim, we used paper-and-pencil tests with pedigree problems as open-ended items (Wools, 2018). In these tests, students were asked to identify the present mode of inheritance and to write down their problem-solving approach as coherently as possible for each pedigree problem. At the end of each pedigree problem, students had to mark the correct mode of inheritance. To visualize these problems, we used two pedigrees (Figure 1), which were generated semi-automatically based on matrices (Surmann, 2017) and plotted with the R package kinship2 (Therneau and Sinnwell, 2015). We selected autosomal pedigrees in which the pattern of inheritance can be determined without any doubt, as other pedigree problems may be too complex for students with little experience in genetics. For precise communication, each person in the pedigree was labeled with a unique identification number, and students were requested to always indicate these numbers when referring to specific family constellations. In the paper-and-pencil tests, each person in the pedigree was also given a name to prevent subjects from confusing the symbols for male and female; this additional information, however, has been omitted here to keep the pedigrees minimalistic.

The trait mapped in the first pedigree (Figure 1a) is inherited in an autosomal recessive pattern. The mode of inheritance can be determined as autosomal recessive based on a single family constellation. Unaffected parents (no. 107 and no. 108) have an affected daughter (no. 113). Therefore, the trait can neither be autosomal dominant nor X-linked dominant, because at least one parent would have to be affected too. In addition, inheritance cannot be X-linked recessive, as the daughter would have inherited an allele from each parent. However, in this case, the father would have been affected too, because men are hemizygous (have only one allele). The pedigree shows inheritance across three generations and includes 14 persons in total, with three females and one male being affected.

The second pedigree problem involves an autosomal dominantly inherited trait (Figure 1b). This pedigree comprises six persons spread over three generations; three of the individuals (two males and one female) are affected. In this case, two family constellations are needed to determine that the inheritance is autosomal dominant. That two affected parents have an unaffected son rules out both recessive modes of inheritance, because parents no. 103 and no. 104 could not have possessed or transmitted any allele that is not associated with the trait, and so their son (no. 111) should have been affected as well. At the same time, an X-linked dominant inheritance is ruled out by the constellation of the unaffected daughter (no. 105) and the affected father (no. 101): Because men have only one X chromosome, the daughter would have inherited the relevant allele in any case and should, therefore, have been affected too.

Sample

For this study, we analyzed paper-and-pencil tests of 89 secondary students. In German secondary schools, pedigree problems are generally part of a basic genetics unit as well as an advanced genetics unit. The advanced unit takes place at the upper secondary level, usually 2 years after the basic unit. Although pedigree problems are thus part of both units, they are treated in much more detail in the advanced genetics unit. To cover a maximum of the variance expected in secondary school, we decided to include students who have only basic experience with pedigree problems as well as students with advanced experience. We examined 45 secondary students (M_{Age} = 14.9), who completed a basic genetics unit, and 44 secondary students (M_{Age} = 16.8) who had participated in the advanced genetics unit. Although we do not possess detailed information about the scope and process of the genetics units in the participating classes, pedigree problems are typical tasks covering scientific inquiry in genetics. For instance, pedigree analysis can be covered by an exemplary task within the educational standards for middle school graduation, which define an expected average achievement level for students at the end of lower secondary education (Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany, 2005). In that task, students are expected to be able to determine the mode of inheritance in the presented pedigree by, among other things, ruling out an X-linked inheritance. Furthermore, pedigree problems are a typical task for the biology exams in the Abitur, the final examination at the end of the higher secondary level. According to the curriculum, upper secondary students are expected to formulate hypotheses on X-linked and autosomal modes of inheritance and justify the hypotheses with available data based on meiosis (Ministry for School and Further Education of the State of North Rhine-Westphalia, 2014, p. 32). Students taking an advanced course (in contrast to those who take a basic course) should additionally be able to consider two-factor analysis, genetic linkage, and crossing-over thereby. As a result, students from both groups should be familiar with all four modes of inheritance and with pedigree analysis in principle. Nevertheless, it is reasonable to assume that the students who participated in the advanced course have more differentiated knowledge and more procedural skills, because they had already attended the basic genetics unit a few years earlier. The advanced course can therefore build on these basics. Following ethical and legal guidelines, all participants were informed about the aim and procedure of the study in advance and participated voluntarily. The results were not used for evaluation or grading. Only necessary data were collected, and they were treated in accordance with the local data-protection laws. As the study was conducted in Germany, students answered in German; quotations in this article were translated into English by the authors.

Coding and Analysis

To answer our research question, we performed a content analysis of students' reasoning on pedigree problems using MAXQDA 2018 (VERBI Software, 2018). For this purpose, the students' written answers for the open-ended tasks were transcribed. We defined propositions as coding units to analyze students' reasoning at the level of individual conclusions (Hackling and Lawrence, 1988). We developed a codebook based on theory and data as a guide for analyzing the students' answers. Therefore, the codebook (Table 1) contained theory-driven and data-driven codes (DeCuir-Gunby *et al.*, 2011), which we present in detail here.

First, we categorized whether the proposition was at all a conclusion or what type the proposition was. In this regard, we distinguished five types of propositions—"proof," "clues," "allegations," "descriptions," "other"—based on the present transcripts but strongly aligned with the theoretical frameworks of Toulmin (2003) and Smith (1988). Based on the classification of Toulmin (2003), a full argument contains at least a claim based on data. In pedigree analysis, typical arguments formulate claims regarding the confirmation or exclusion of at least one type or mode of inheritance based on data from the pedigree as evidence; for example, "the mode of inheritance is autosomal [claim], as both sexes are affected [data]" (Student 1009|Line 40). Therefore, arguments regarding the mode of inheritance represent condition–action pairs, as described by Smith (1988). However, these arguments can also be further classified from a scientific perspective. Arguments regarding the mode of inheritance can represent either clear-cut decisions or indications for or against individual modes of inheritance. For example, in the case of X-linked recessive inheritance, more males tend to be affected than females. If such a distribution is present, this can therefore be an indication of an X-linked recessive inheritance, but it is not suitable as definitive proof. To account for this distinction, we have used two categories: proof and clues. In proof-type arguments, one mode of inheritance is unambiguously confirmed or ruled out, whereas arguments that provide evidence for or against modes of inheritance without making a clear decision were called clues. Although

TABLE 1. Variables and categories used to characterize students' reasoning on pedigree problems (a complete version including code descriptions and examples can be found in the Supplemental Table 1)

Variable	Categories
Type of statement	Description
	Allegation
	Clue
	Proof
	Other
Pedigree feature	Phenotypic family constellation
	Genotypic family constellation
	Proportion of affected persons
	Gender ratio among affected
	Distribution of affected over generations
	Pedigree size or structure
	Other
	Missing
Claim	Autosomal dominant
	Autosomal recessive
	X-linked dominant
	X-linked recessive
	Autosomal
	X-linked
	Y-linked
	Dominant
	Recessive
	Other
Orientation	Confirmation
	Refutation
Conclusiveness	Conclusive
	Indeterminable
	Inconclusive

technical in nature, this distinction can also be understood in terms of Toulmin's (2003) structure. Arguments that provide evidence for or against single heritages but do not allow a clear decision usually contain what Toulmin (2003) calls a qualifier. This is an optional argument component that indicates some uncertainty, as we can see in the following statement: "Again, both sexes are affected [data], which again indicates [qualifier] autosomal inheritance [claim]" (1012|41). Beyond these two central categories, three other types of statements were distinguished that do not constitute complete arguments according to Toulmin (2003). In this regard, two categories—"allegations" and "descriptions"—represent incomplete arguments, because they consist of only one element of Toulmin's (2003) classification each. Allegations represent stand-alone claims that do not contain any data as evidence. Descriptions mention data from the pedigree but do not contain a claim. Finally, statements that did not fit into one of the first four categories were classified as "other."

To differentiate the various types of pedigree features students cited as data in their arguments (or descriptions), we distinguish eight categories based on students' statements and results from previous research. A typical approach during pedigree analysis is to test multiple hypotheses based on informative family constellations or by assigning genotypes (Hackling, 1994). Based on the written arguments, we cannot determine whether students identified a family constellation by recalling memorized production rules or assigning genotypes. Nevertheless, we were able to distinguish whether they argued using genotypes or phenotypes. In cases in which students mentioned a particular family constellation in a statement, we classified the statements according to whether the students were referring to the phenotypic level (coded as "phenotypic family constellation") or genotypic level (coded as "genotypic family constellation"). Another category we found in our data regarding pedigree analysis includes statements referring to the "proportion of affected persons" in relation to the total number of persons. This category addresses the misconception that dominance and allele or trait frequency are linked, which has been described in previous work, not just in the context of pedigree analysis (e.g., Smith and Good, 1984; Smith, 1988; Abraham *et al.*, 2014). This becomes evident in the following example: "Since very few people are affected, the inheritance is recessive" (1003|34). The starting point for another pair of categories was a case study presented by Hackling (1994), in which one subject determined the mode of inheritance to be autosomal dominant because there were affected individuals in each generation (mode of inheritance cannot be recessive, must be dominant) and males and females were affected (mode of inheritance cannot be X-linked). Based on the intersection of Hackling's (1994) descriptions and our data, we created categories to capture the cases in which students quote the "gender ratio among affected" individuals or the "distribution of the affected across generations." Not captured by the categories mentioned so far is the fact that students also referred to basic features of the pedigree, such as the number of generations represented. Therefore, we added a category to cover statements in which students referred to the overall "pedigree size or structure." In addition, sometimes a datum from the pedigree was "missing" or students mentioned "other" pedigree features we were unable to categorize more accurately.

To analyze students' claims more precisely, we introduced categories based on content knowledge and added all relevant modes of inheritance ("autosomal dominant," "autosomal recessive," "X-linked dominant," "X-linked recessive," or "Y-linked") as well as more general conclusions on heredity ("dominant," "recessive," "autosomal," or "X-linked") as categories. Here too, claims that could not be assigned to any of these categories were coded as "other." While the previous two variables account for the biology-specific content of the propositions, we also analyzed students' strategic procedures. The extent to which the students used a falsification strategy (Smith, 1988) could be deduced by the orientation of their arguments. In this regard, we distinguished statements that entailed support for or confirmation of claims from those weakening or rejecting claims. As a result, we categorized each argument as "confirmation" or "refutation" with regard to the subject of the claim. The goal here was to investigate the extent to which students used a falsification strategy in their pedigree analyses. Finally, we classified each argument in terms of its conclusiveness, as either "conclusive" or "inconclusive" or, if necessary, as "indeterminable" based on expert knowledge.

To ensure the consistency of the coding, we selected a random set of 26.7% of all propositions, which were coded by two

independent raters (JT and MG, a student assistent; O'Connor and Joffe, 2020). We calculated Cohen's kappa (Cohen, 1960) for all main categories of interest (Table 1) and interpreted the values in the context of the limits described by Landis and Koch (1977). We interpreted these values to check for inconsistencies and revised the coding manual accordingly (O'Connor and Joffe, 2020). Then, the coding process was repeated until satisfactory agreement was obtained for all categories. According to the thresholds described by Landis and Koch (1977), we achieved almost perfect agreement ($0.8 < \kappa \leq 1$) for the pedigree feature ($\kappa = 0.81$), claim ($\kappa = 0.92$), and argument orientation ($\kappa = 0.98$). Substantial agreement ($0.6 < \kappa \leq 0.8$) was achieved for type of statement ($\kappa = 0.75$) and argument conclusiveness ($\kappa = 0.77$). We considered these values to be satisfactory, and so one rater (MG) coded all remaining propositions to obtain the full set of data used in this study.

To analyze the coding results, we created tables to examine what kinds of propositions and which claims the students expressed for each pedigree. To examine the connections between students' claims and the associated data, we created a mosaic plot. Mosaic plots are used to visualize contingency tables and consist of tiles that represent the individual frequencies through tile size (Hofmann, 2008).

In addition, we determined which pedigree members students mentioned in their reasoning by calculating frequencies for all pedigree members' identification numbers. To visualize these data, we drew the individual frequencies directly into the pedigrees. To obtain comprehensible diagrams, we applied the principles of heat maps and bubble charts to present our data. This means that the relative frequencies are represented by the size and color of the bubble drawn on top of each family member's symbol. A large red bubble indicates that a person has been mentioned frequently, while a small yellow bubble indicates that the person has been mentioned very rarely.

RESULTS

The total data set consists of 516 propositions, on average 6.1 per student (excluding five students who did not answer at all). Overall, the number of propositions made was slightly higher for the first task (the autosomal recessive pedigree) than for the second task (the autosomal dominant pedigree). The proportion of statements representing full arguments (clues or proof) was considerably higher in the group of advanced genetics students, $\chi^2(1) = 36.68$, $p < 0.001$ (see Table 2 for detailed results).

Beginning Genetics Students Do Not Use a Falsification Strategy

To examine whether students used a falsification strategy to identify the mode of inheritance by ruling out all alternative modes, we analyzed their arguments in multiple regards. Beginning genetics students predominantly (92.6%) formulated confirmatory statements; for example: "Both men (103) and women (104) have the trait, so it is autosomal" (1023|41). Advanced genetics students, in contrast, formulated confirming (58.9%) as well as falsifying arguments (41.1%); for example: "Gonosomal dominant can be excluded on individuals 101, 102, and 104" (B15|41). This difference is significant, $\chi^2(1) = 40.24$, $p < 0.001$, and demonstrates that the advanced genetics students resorted to a falsification strategy at least temporarily, while the beginning genetics students made a straightforward decision on the mode of inheritance. However, these two quotes also show a difference in scale. Arguments that refer to a specific mode of inheritance—for example, autosomal dominant or X-linked recessive inheritance—were rare in the group of beginning genetics students but rather common in the group of advanced genetics students (see Table 3). In both groups, autosomal recessive inheritance was the most frequently mentioned specific mode of inheritance. In addition, arguments that either indicated whether inheritance was sex-linked or drew a conclusion about the trait's dominance without naming a specific inheritance mode were common in both groups (see Table 3). To examine the quality of an argument, however, we found it necessary to look at the data used to substantiate the claims.

A Variety of Decisions Are Based on Superficial Pedigree Features

Students referenced a variety of pedigree features as data to prove their claims; for example, the gender ratio among affected persons, the proportion of affected persons, or the distribution of the trait over generations. We call these features "superficial," because they are easily noticed but allow no or only vague conclusions (clues). Family constellations, either at the phenotypic or genotypic level, form the counterpart. They must be remembered or analyzed precisely. Unlike "superficial" features, family constellations allow clear decisions. In fact, many students described specific family constellations, often at the phenotypic level, sometimes at the genotypic level. In total, 89.9% of all students referred at least once to a specific family constellation. However, not every proposition contains a conclusion. To identify which pedigree features were mentioned in connection

TABLE 2. Number and types of propositions for both groups of students[a]

Group	Subjects	Task/pedigree	Propositions	Propositions PP	Type of proposition			
					Descriptions	Allegations	Arguments	Other
Beginning genetics students	39	Autosomal recessive inheritance	134	3.44	64 (47.8%)	14 (10.4%)	48 (35.8%)	8 (6%)
	40	Autosomal dominant inheritance	113	2.83	46 (40.7%)	16 (14.2%)	44 (38.9%)	7 (6.2%)
Advanced genetics students	43	Autosomal recessive inheritance	153	3.56	34 (22.2%)	9 (5.9%)	93 (60.8%)	17 (11.1%)
	43	Autosomal dominant inheritance	116	2.70	18 (15.5%)	8 (6.9%)	80 (69%)	10 (8.6%)

[a]Students who did not respond to a task were excluded from the individual calculation. Propositions PP: Propositions per person.

TABLE 3. Subjects students refer to in their conclusions by task and group[a]

				The claim refers to …							
Group	Subjects	Task/pedigree	Propositions	Dominant inheritance	Sex-linked Inheritance	AD	AR	XD	XR	Y	Other
Beginning genetics students	23	Autosomal recessive inheritance	48	26 (54.2%)	10 (20.8%)	0 (0%)	2 (4.2%)	1 (2.1%)	2 (4.2%)	0 (0%)	7 (14.6%)
	26	Autosomal dominant inheritance	44	23 (52.3%)	12 (27.3%)	1 (2.3%)	4 (9.1%)	0 (0%)	0 (0%)	0 (0%)	4 (9.1%)
Advanced genetics students	39	Autosomal recessive inheritance	93	34 (36.6%)	16 (17.2%)	5 (5.4%)	11 (11.8%)	6 (6.5%)	9 (9.7%)	1 (1.1%)	11 (11.8%)
	37	Autosomal dominant inheritance	80	31 (38.8%)	16 (20%)	7 (8.8%)	6 (7.5%)	9 (11.2%)	4 (5%)	0 (0%)	7 (8.8%)

[a]Students who did not respond to a task or did not make any analytical statements there were excluded from the individual calculation. AD, autosomal dominant inheritance; AR, autosomal recessive inheritance; XD, X-linked dominant inheritance; XR, X-linked recessive inheritance; Y, Y-linked inheritance.

with which type of inheritance, we analyzed the relationship between claim (on the mode of inheritance) and data mentioned (the pedigree feature to justify that claim; Figure 2).

It becomes apparent through analysis of the mosaic plots (Figure 2) that advanced genetics students based their decisions regarding the mode of inheritance primarily on specific family constellations within the pedigree. The proportion of statements referring to family constellations was significantly larger in the group of advanced genetics students (75.5%) compared with the group of beginning genetics students (46.9%), $\chi^2(1) = 18.02$, $p < 0.001$. In addition, a significantly larger proportion of advanced genetics students' arguments focused on specific modes of inheritance (37.4% compared with 12.3% of basic genetics students' arguments), $\chi^2(1) = 15.11$, $p < 0.001$. The distribution of affected over generations and the gender ratio among affected individuals were cited fairly often, each in a specific context. The distribution of affected over generations was primarily used to prove the trait's dominance (12 of 15 mentions were associated with this kind of claim), and the gender ratio among affected individuals was primarily mentioned in connection with sex linkage. Indeed, most arguments by beginning genetics students regarding sex linkage were based on the gender ratio among affected, while even references to specific family constellations were rare in this case. Compared with advanced genetics students, beginning genetics students referred significantly more often to the gender ratio among affected in their arguments on sex linkage, $\chi^2(1) = 4.69$, $p = 0.030$. Overall, beginning genetics students argued more frequently on the basis of superficial pedigree features, $\chi^2(1) = 21.18$, $p < 0.001$, and less frequently on the basis of genotypes, $\chi^2(1) = 7.84$, $p = 0.005$. A rare kind of data, primarily used by beginning genetics students (eight out of 10 mentions occurred in this group) was the proportion of those affected. It was cited primarily in relation to the trait's dominance (seven of the total 10), which, however, is not appropriate, because trait prevalence depends on allelic frequency alone and does not allow any indications of dominance. Accordingly, arguments based on this kind of evidence are inconclusive in any case. The conclusiveness of arguments relating to specific family constellations, in contrast, cannot be assessed across the board. An extended mosaic plot showing which arguments were classified as correct and at what frequency can be found in the Supplemental Material. An argument does not need to be conclusive just because a particular constellation in the pedigree is pointed out. To analyze in detail which family members the students referred to, we used the identification numbers of the pedigree members. The frequency of the mentions was then drawn in the pedigree, separately for each task and student group (Figure 3).

The first pedigree (Figure 3a and b) shows the inheritance of an autosomal recessive trait. In this case, all alternative modes of inheritance can be excluded by one family constellation (unaffected father no. 107 and mother no. 108 in combination with their affected daughter no. 113). In the group of beginning genetics students, however, the affected but uninformative person no. 104 was mentioned most often in arguments regarding inheritance (14 mentions; relative frequency: 100%). Also very frequently mentioned, in descending order, were the other affected no. 113 (93%), no. 105 (86%), no. 102, and no. 107 (each 64%). Reasonably often cited in arguments were person no. 108 (50%), the children of no. 103 (29%) and no. 104 (no. 110, no. 111, and no. 112; each 43%). In contrast, the picture is markedly different for advanced genetics students' mentions. Most frequently mentioned was the informative person no. 113 (39 mentions; relative frequency: 100%) closely followed by the parents no. 107 (95%) and no. 108 (92%). The other affected family members, in contrast, are cited notably less frequently. Their mentions in descending order: no. 104 (56%), no. 102 (46%), and no. 105 (21%). Thus, the affected person no. 105 is mentioned less frequently than no. 101 (28%) and the children of no. 103 (also 21%) and no. 104 (no. 110: 26%; no. 111: 28%; no. 112: 23%). The relative frequency of mentions of the unaffected and uninformative individuals no. 106, no. 114, and no. 115 is low in both groups and does not exceed 15%. Overall, comparing the arguments referring to family constellations in terms of group membership reveals that the importance of the family constellation consisting of family members no. 107, no. 108, and no. 113 seems rarely to be recognized by the group of beginning genetics students. The advanced genetics students, in contrast, seem to recognize the importance of this family constellation.

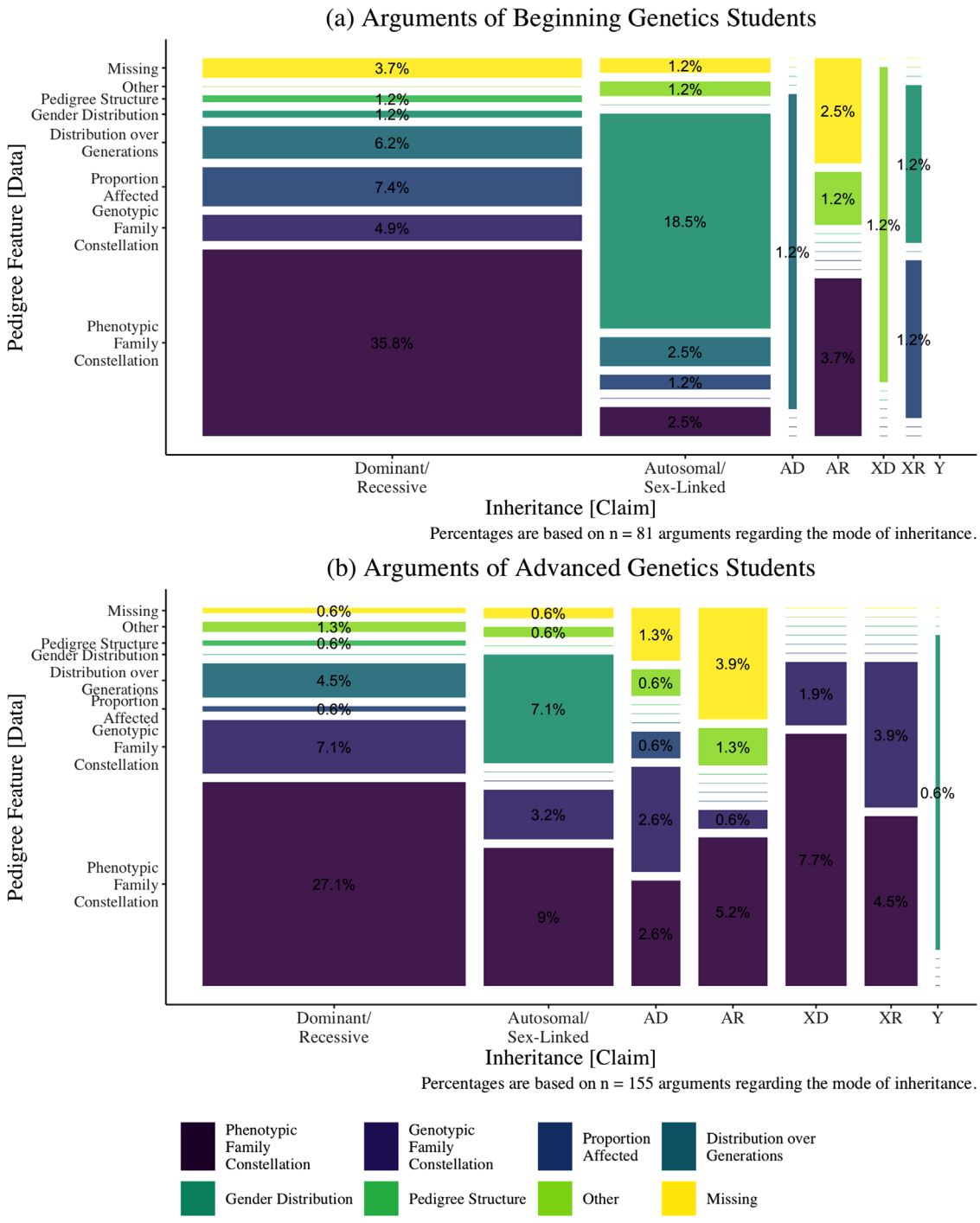

FIGURE 2. Mosaic plot showing the structure of students' arguments. The relationship between the inheritance mentioned in the argument and the data used to prove that claim is shown separately for each group. The diagram maps all possible combinations of pedigree features and claims regarding the inheritance, while the area of each box represents how often that specific combination was found among students' statements. Thin lines without percentages indicate that a specific combination was not found. AD, autosomal dominant inheritance; AR, autosomal recessive inheritance; XD, X-linked dominant inheritance; XR, X-linked recessive inheritance; Y, Y-linked inheritance.

The second pedigree (Figure 3c and d) shows the inheritance of an autosomal dominant trait. For the constellation given, the exclusion of all alternative modes of inheritance is possible based on two individual constellations. Based on the family constellation of the affected parents no. 103 and no. 104 in combination with their unaffected child no. 111, both recessive modes of inheritance are ruled out. An X-linked dominant inheritance can be excluded, because no. 105 is not

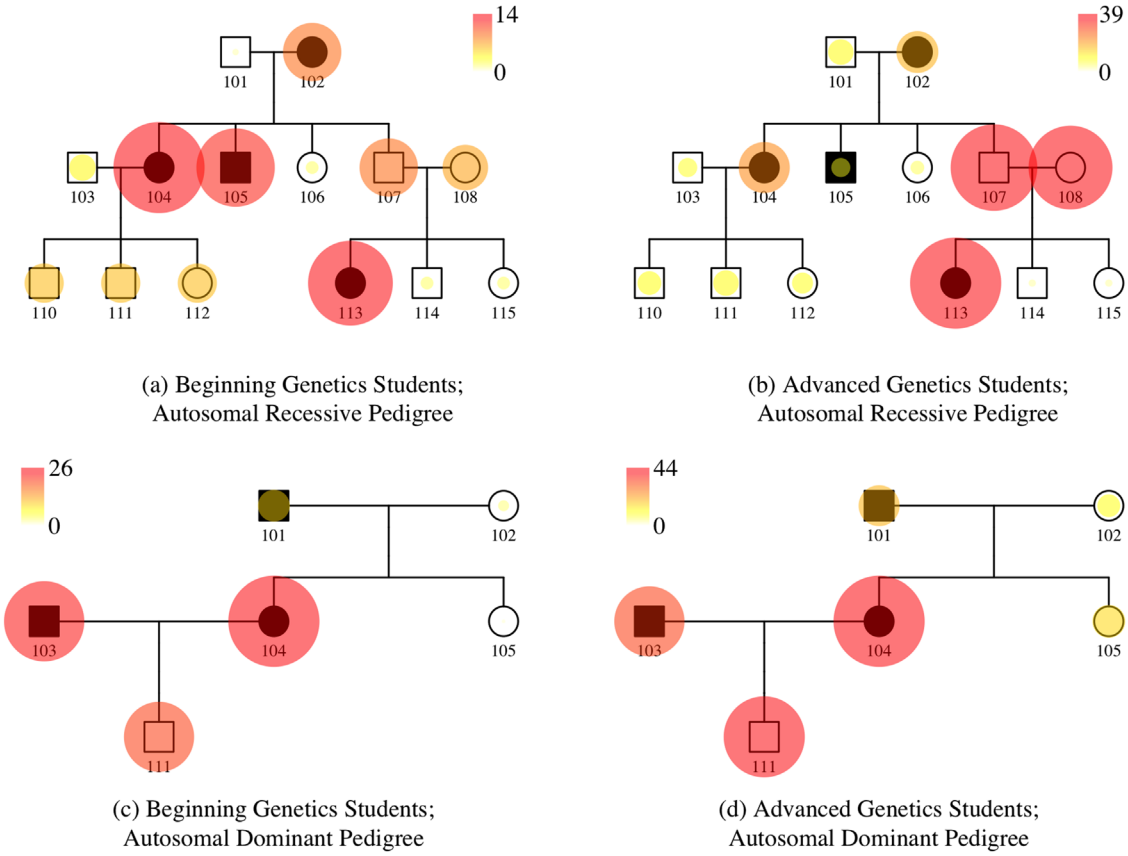

FIGURE 3. Representation of the frequency with which students mentioned individual family members in their arguments regarding the present mode of inheritance separately for each pedigree and group. The frequencies are shown in the pedigrees by the color and size of the points. Mentions of individual family members in the pedigree showing an autosomal recessive trait by (a) beginning genetics students and (b) advanced genetics students. Mentions of individual family members in the pedigree showing an autosomal dominant trait by (c) beginning genetics students and (d) advanced genetics students.

affected but should have inherited the trait causing allele from father no. 101. Beginning genetics students refer in their arguments regarding the mode of inheritance mainly to the family constellation of mother no. 104 (26 mentions; relative frequency: 100%), father no. 103 (88%), and their son no. 111 (77%). Beyond that, only the affected male founder no. 101 is mentioned with noteworthy frequency (35%). The female founder no. 102 is cited rarely (12%), while the informative female no. 105 is mentioned only once (4%). Comparing these results with those of the advanced genetics students, it is noticeable that differences can be observed especially among those family members not mentioned by the beginning genetics students. In this group too, the informative family constellation of mother no. 104 (44 mentions; relative frequency 100%), father no. 103 (77%), and their son no. 111 (89%) is mentioned by far most frequently. However, the informative female no. 105 (36%) and her father no. 101 (45%) are also mentioned moderately often. In this group, least often mentioned is the female founder no. 102 (25%). These results indicate that both groups particularly recognize the family constellation of family members no. 103, no. 104, and no. 111 as relevant and formulate arguments based on this constellation. The second informative family constellation (no. 101, no. 102, and no. 105), necessary to rule out an X-linked dominant inheritance, is hardly ever mentioned by beginning genetics students. Advanced genetics students, in contrast, create arguments based on this family constellation too.

At the end of each task, students had to mark in a multiple-choice item which inheritance they thought was present. Overall, the beginning genetics students were less likely to answer correctly than the advanced genetics students. In the case of the autosomal recessive pedigree, the majority of students of both groups opted for the correct mode of inheritance (beginning genetics students: 65.9%; advanced genetics students: 88.1%). For the autosomal dominant pedigree, however, a majority of beginning genetics students again chose autosomal recessive inheritance (55.8%; only 32.6% decided correctly for autosomal dominant inheritance). Of the advanced genetics students, 56.1% correctly selected the autosomal dominant mode of inheritance. However, a correct decision in the multiple-choice item does not necessarily indicate that the decision is explained based on correct arguments. It can be stated that, in most cases, students chose the correct mode of inheritance if at least one of their arguments was found to be conclusive (see Figure 4). Even students who had not formulated any conclusive arguments in a pedigree problem sometimes identified the correct inheritance. The proportion is even quite large in the case of autosomal recessive inheritance, especially among

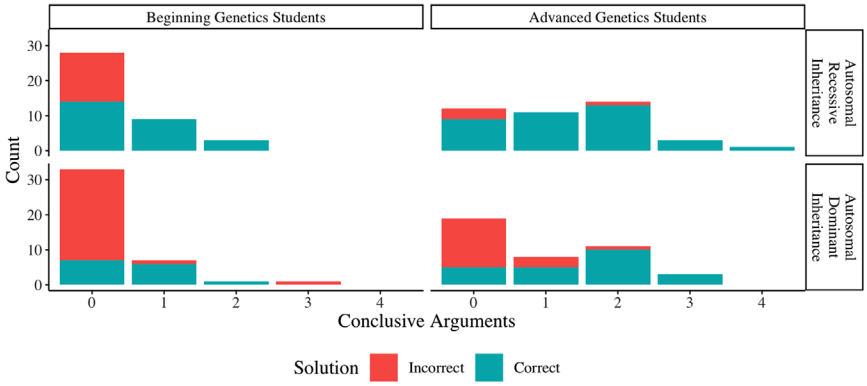

FIGURE 4. Number of conclusive arguments per person and pedigree. Individuals were counted as a function of group membership. Colors indicate which proportion of individuals marked the correct inheritance in an attached multiple-choice item.

advanced genetics students (see Figure 4). For an in-depth analysis of which arguments were rated as correct in which group, see Supplemental Figure 1.

Students' Use of Unreliable Arguments

As the previous results suggest, analytical statements made by advanced genetics students (60.7%) were more frequently conclusive compared with statements by beginning genetics students (29.3%), $\chi^2(1) = 22.37$, $p < 0.001$. In this section, we will analyze some of these inconclusive arguments in detail. We start by presenting some arguments in which students decide between dominant and recessive inheritance and then provide some arguments regarding sex linkage.

A typical conclusion based on the proportion of affected persons among all pedigree members was made by one student, who stated: "Since very few people are affected, the inheritance is recessive" (1003|34). In this statement, Student 1003 implies that autosomal dominant inheritance is indicated by the fact that a large number of people are affected. The proportion of people affected is highly dependent on chance and strongly influenced by the frequency of the trait-determining allele in the total population. Chance plays an important role, because pedigrees are usually too small for probability-based evidence. A carrier of an autosomal dominant trait may pass the trait to all his children or none; starting from one allele, all descendants could be affected or none of them. Nevertheless, the allelic frequency has an impact on how likely an autosomal recessive inheritance is (Hackling and Lawrence, 1988). If the autosomal recessive trait is very common in the population, it is very likely that a considerable number of the family members who are marrying in are carriers. This would mean, in turn, that a large proportion of the pedigree members could be carriers. If an allele is very rare and many pedigree members are affected, an autosomal recessive inheritance is in fact rather unlikely, as it would require several unrelated persons marrying in to bear the rare allele. Because there was no information on allelic frequency for either pedigree problem here, the proportion of persons who are affected does not allow any conclusions to be drawn about the inheritance mode.

A typical conclusion based on trait distribution over generations is made by Subject B6: "First of all, it must be noted that the pedigree cannot be dominant, since the trait does not occur in every generation" (B6|40). What B6 observes here is the distribution of the trait over generations, which indeed can be used to rule out simple dominant inheritance (under the assumptions of full penetrance and no de novo mutations). A correct rule could state that a skipped generation rules out dominant inheritance, because that is about non-affected parents having an affected child. The comparison makes clear that the rule presented by Subject B6 deviates slightly from the appropriate rule and does not necessarily lead to correct results. For example, no statement can be made on the basis of a terminal generation without affected persons. While the slightly modified rule results in the correct decision for the autosomal recessive pedigree, Subject B6 de facto considers a dominant inheritance impossible for the autosomal dominant pedigree as well based on his rule. In return, his rule implies that a dominant trait is always transmitted. This is a recurring misconception, appearing, for example, in a statement by Subject B20: "The inheritance is also recessive, since none of the children [110, 111, 112] of 103 and 104 carry the trait. If the inheritance was dominant, 110, 111 or/and 112 would carry the trait, since the mother 104 carries the trait" (B20|41–42). In this statement, the student assumes that the trait cannot be dominant, because otherwise an affected mother would certainly have passed it on to at least one child. What we cannot say is *which* false assumption underlies this statement: On the one hand, the decision could be based on the expectation that a dominant trait will always necessarily be passed on to the next generation, which is not necessarily the case; on the other hand, the decision could be based on an incorrect understanding of probabilities, by assuming that, in the case of three children, at least one should be affected.

Based on the gender ratio, there were two different ways of reasoning evident. On the one hand, the distribution of the trait by gender was used as a qualitative measure: "The trait is inherited as an autosomal trait since both female and male persons are affected" (D11|40). In this argument, it seems irrelevant how many people of each gender are affected, as long as there is at least one each. On the other hand, the distribution of the trait depending on gender was used as a quantitative measure, too: "[T]he inheritance is gonosomal because almost only women are affected" (1003|36). In this case, an (extreme) imbalance was used as an argument for sex linkage.

DISCUSSION

In this study, we wanted to explore secondary school students' approaches to solving pedigree problems and how they constructed arguments to support their decisions. For this purpose, we analyzed two pedigree problems representing autosomal modes of inheritance that should be known to all students. Overall, we identified a number of interesting findings.

While the students' answers hardly differed in number of propositions by group, beginning genetics students generated significantly fewer arguments. A notable proportion of this

group actually stated no arguments at all. In fact, many of the beginning genetics students' statements were descriptive. Nevertheless, a considerable number of students correctly identified the autosomal recessive inheritance in the first pedigree, even in the absence of conclusive arguments (see Figure 4). There could be several reasons why the students' answers turned out the way they did. First, some students may have found it difficult to write down their approaches. Among other things, a lack of expertise could have been responsible for this, because argumentation quality is presumed to be related to content knowledge (Sadler and Zeidler, 2005; Sadler and Fowler, 2006; McNeill *et al.*, 2006; Aufschnaiter *et al.*, 2008). Second, some students may in fact have reached a conclusion on the mode of inheritance without citing any evidence or data. This would be in line with the proposed learning progressions on scientific argumentation, which assume that it is easier to formulate claims than to support them with evidence or data (Songer *et al.*, 2009; Berland and McNeill, 2010; Lee *et al.*, 2014). In fact, this would also fit the observation that a substantial portion of students succeeded in correctly identifying the present mode of inheritance without having formulated any conclusive arguments. Third, some students may have considered a pedigree feature relevant but were unable to draw any conclusion from it. In this case, too, a lack of content knowledge could be the issue. Smith (1988) observed that individuals who were unsuccessful in pedigree analysis tended to quote affected and unaffected descendants of matings shown in the pedigree without drawing conclusions. He stated that "these individuals appear to be looking for recognizable patterns or other cues, but they either do not recognize significant cues or they do not know which of their patterns are meaningful" (Smith, 1988, pp. 416–417). Finally, some students could have assumed that a brief description of the pedigree was part of a conclusive pedigree analysis. However, beginning genetics students not only made more descriptive statements but also referred less frequently to specific family constellations and considered superficial pedigree features as data more often. For example, the numbers of affected and unaffected persons and the ratio between them were repeatedly mentioned, as was the number of generations. In particular, when deciding between autosomal or X-linked inheritance, beginning genetics students mainly relied on superficial features as evidence in their arguments. This is in line with the results of Hackling (1994), who found that the majority of novice undergraduates lack the necessary content knowledge to reliably check for X-linked inheritance. Although the advanced genetics students referred to family constellations much more frequently, the only informative family constellation to rule out X-chromosomal dominant inheritance in pedigree 2 was almost never mentioned by beginning genetics students and only sometimes by advanced genetics students. In general, beginning genetics students focused primarily on affected family members whether or not they were informative; thus, a considerable proportion of students in this group did not seem to be able to distinguish between informative and uninformative family constellations. Advanced genetics students, in contrast, were particularly concerned with informative pedigree members, which indicates that they identified critical cues more often. This was true at least in the case of the autosomal recessive pedigree; for the autosomal dominant pedigree, however, the results were much less obvious. Therefore, a clear trend cannot be proven empirically or based on previous studies. Hackling and Lawrence (1988) observed that "experts recognized more critical cues than the students" (p. 537). In the study by Smith (1988), however, unsuccessful and successful participants identified comparable numbers of critical cues, whereas unsuccessful participants identified noticeably more noncritical cues.

Nevertheless, identifying a critical cue does not necessarily mean drawing the right conclusions from it. This is independent of whether the cue was identified based on assigning genotypes or by recalling a memorized production rule. In both cases, genetic knowledge is required. When a certain condition occurs in the pedigree, for example, when two unaffected parents have an affected child, the corresponding action that a dominant inheritance is ruled out comes to pass. In our sample, only about every fourth argument by the beginning genetics students was technically sound, and even among advanced genetics students, more than one in three conclusions was incorrect. This observation is largely consistent with Smith's (1988) comment that unsuccessful students often either do not use production rules at all or use them inaccurately, incompletely, or even incorrectly. Indeed, beginning genetics students often used superficial and uninformative pedigree features as data, which was probably the reason for the low number of valid conclusions.

Regarding the strategic approach, we highlight two points. First, advanced genetics students much more often used a falsifying procedure (cf. Hackling and Lawrence, 1988), whereas beginning genetics students almost exclusively generated confirmatory arguments. Second, but just as important, students in both groups primarily did not test the modes of inheritance individually, but rather decided between dominant and recessive, on the one hand, and autosomal and X-linkage, on the other hand.

Limitations

The analysis of pedigree problems is time-consuming, and all results described here were derived from the analysis of only one pedigree per type of autosomal inheritance. Based on the available data and results, it is therefore not possible to make any predictions about the general strategic approach of secondary students to other pedigrees, especially those that represent X-linked inheritance. Apparently, it is uncertain whether individual pedigree features influence students' approaches and, if so, how. It is difficult to imagine varying multiple pedigree features systematically in order to investigate the effects of these variations on students' procedure. The influence of features such as pedigree size or the proportion of affected can only be investigated, if at all, in large samples in which all subjects analyze only a subset of the pedigrees.

In our study, all students participated voluntarily, and their results were neither evaluated nor graded. Accordingly, it is reasonable to assume that some individuals may not have put full effort into analyzing both pedigrees and writing down their approaches. Nevertheless, we are convinced that we were able to uncover students' typical patterns of reasoning and to identify what kind of evidence students typically use to justify their decisions. We examined a reasonably large sample of students in order to quantify our main results. However, this had an impact on how precisely we were able to analyze each student's

data. Because we analyzed written responses, it is likely that we missed some of the students' thoughts. For example, we assume that they may not have written it down if they recognized a constellation as meaningful but were unable to draw any conclusions from it. To uncover such cases, one would have to use other methods that provide more insight into the thinking processes of the students, such as retrospective interviews, thinking aloud, or eye tracking (Lai *et al.*, 2013). Because retrospective interviews and think-aloud protocols could be potentially biased by mental processes, eye-tracking data could be even better, because information intake is recorded directly and in an unfiltered manner. For this, however, the sample size would have had to be reduced significantly.

Implications for Teaching

Pedigrees are an excellent way to learn genetics, for multiple reasons, and are a core practice in genetic classes. Pedigree problems represent realistic scientific problems and illustrate real human genetic phenomena observable at the macroscopic level (Tsui and Treagust, 2010). They are well suited for scientific reasoning and are appropriate to help students develop a comprehensive understanding of meiosis and Mendelian inheritance in humans.

Our findings offer some implications for what students should consider when analyzing pedigrees. First, a significant proportion of the participants' conclusions were based on simplified rules, which may serve as a first indication but are not suitable for making clear decisions. We suppose these students were relying too much on their instincts. Students need to learn how to distinguish a presumption based on superficial features from a decision based on clear evidence. To achieve this, it is advisable to also examine pedigrees where the first impression is misleading; for example, a recessive pedigree in which most people are affected might be a good start. Second, we suppose that it is not necessary to start by memorizing all informative family constellations. Students should rather learn to analyze family constellations on demand, for instance, by assigning genotypes and using their knowledge of meiosis to test whether there is a contradiction for one mode of inheritance. Of course, this requires students to be able to correctly assign genotypes, even for sex-linked modes of inheritance. In a survey by Hackling (1994), however, a large number of novice undergraduate university students were unable to assign genotypes for X-linked dominant inheritance accurately. We suppose that it would be worth considering practicing this method before considering X-linkage. Beginning genetics students made almost exclusively confirmatory statements, whereas a complete and conclusive procedure is usually characterized by the fact that all alternative inheritance modes are conclusively ruled out by testing multiple hypotheses (Hackling and Lawrence, 1988). To encourage this approach, the task could be changed for practice, prompting students to explicitly exclude impossible inheritance modes one by one. Only when the procedure has been practiced should X-linkage be introduced as a modifier. Apart from the strategic approach, students should also be supported in setting up coherent arguments consisting of claims that are justified based on data. One way to do this would be to use scaffolding (McNeill *et al.*, 2006).

Tsui and Treagust (2010) highlight that the use of human genetic pedigrees representing Mendelian problems can promote a deterministic notion of genes. To address this issue, both clear and simple pedigrees and ambiguous pedigrees should be used. It is possible to use ambiguous pedigrees (eventually showing complex traits) to generate research questions or hypotheses for further teaching; this would put a special focus on scientific inquiry. It should be possible, for example, to introduce the phenomenon of incomplete penetrance by using a pedigree that does not seem to fit with any known mode of inheritance. Stewart *et al.* (2005) described the use of human pedigrees in relation to blood types and achondroplasia to demonstrate that the model of simple dominance is not always adequate. These observations are the starting point for studying these "anomalous" inheritance patterns (Stewart *et al.*, 2005). Naturally, X-linkage could be considered a modification of simple dominance in the same way (Collins and Stewart, 1989). In particular, pedigrees with an unexpected gender distribution among those affected would be a good starting point here. Overall, by emphasizing reasoning and scientific inquiry, pedigree analysis can make a big contribution to the development of genetic literacy at large.

Teaching and practicing pedigree analysis should not be limited to teaching the use of content-specific concepts. Rather, the falsification strategy should also be taught and practiced. One possible way to do so would be to encourage students to use a falsification strategy with the instructions given. For example: "Try to exclude as many modes of inheritance as possible. Justify your decisions." The use of a falsification approach could be triggered even more vigorously by further structuring and dividing the pedigree analysis into individual tasks per inheritance. For example: "Complete the following sentence. The trait is not inherited in an autosomal recessive manner because …" Moreover, it should be emphasized that a rigorous use of the falsification strategy can help check the results for plausibility and ensure that the analysis is complete. A complete analysis should provide an evaluation for all four modes of inheritance. In pedigree problems used in schools, at least two modes of inheritance can often be ruled out without doubt. Even in the case of an X-linked mode of inheritance, where no definite decision is possible, there are usually recognizable clues in favor of one mode of inheritance.

Because our research, like previous studies on pedigree analysis, did not specifically focus on teaching pedigree analysis, there are still a number of unanswered questions. In pedigree analysis, it is of primary importance to identify critical cues and to draw the right conclusions from them. In our opinion, however, it is still unclear which is the most effective way to teach students to identify these critical cues. On the one hand, there could be an ideal strategy to teach this ability; on the other hand, it could simply develop through practice. In any case, practice certainly plays an important role in pedigree analysis; therefore, students should definitely be offered a variety of pedigrees to practice with (see Timm *et al.*, 2020). Manifold practice opportunities are particularly useful if students can write down arguments for their solutions and receive individual feedback each time. This would be very time-consuming for teachers but could possibly be automated via a machine-learning approach (see Zhai *et al.*, 2021) in which students' arguments are automatically categorized to select appropriate feedback.

ACKNOWLEDGMENTS

We thank Dr. Yvonne Lettmann for good advice and fruitful discussions and Marleen Gerwin for helping with the coding. This project was financially supported by the German Federal Ministry of Education and Research (BMBF) under the reference number 01PL16075. We made use of professional editing services, which was financially supported by the Interdisciplinary Center for Educational Research (IZfB) of the University of Duisburg-Essen, Germany. We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the funders. We confirm that there are no potential conflicts arising from these arrangements.

REFERENCES

Abraham, J. K., Perez, K. E., & Price, R. M. (2014). The Dominance Concept Inventory: A tool for assessing undergraduate student alternative conceptions about dominance in Mendelian and population genetics. *CBE—Life Sciences Education*, *13*, 2, 349–358. https://doi.org/10.1187/cbe.13-08-0160

Aznar, M. M., & Orcajo, T. I. (2005). Solving problems in genetics. *International Journal of Science Education*, *27*(1), 101–121. https://doi.org/10.1080/09500690410001673801

Bennett, R. L. (2010). *The practical guide to the genetic family history* (2nd ed.). Hoboken, NJ: Wiley-Blackwell. https://doi.org/10.1002/9780470568248

Bennett, R. L., Steinhaus, K. A., Uhrich, S. B., & O'Sullivan, C. (1993). The need for developing standardized family pedigree nomenclature. *Journal of Genetic Counseling*, *2*(4), 261–273. https://doi.org/10.1007/BF00961575

Bennett, R. L., Steinhaus, K. A., Uhrich, S. B., O'Sullivan, C. K., Resta, R. G., Lochner-Doyle, D., … & Hamanishi, J. (1995). Recommendations for standardized human pedigree nomenclature. *Journal of Genetic Counseling*, *4*(4), 267–279. https://doi.org/10.1007/BF01408073

Bennett, R. L., Steinhaus French, K., Resta, R. G., & Lochner Doyle, D. (2008). Standardized human pedigree nomenclature: Update and assessment of the recommendations of the National Society of Genetic Counselors. *Journal of Genetic Counseling*, *17*(5), 424–433. https://doi.org/10.1007/s10897-008-9169-9

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*(5), 765–793. https://doi.org/10.1002/sce.20402

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Collins, A., & Stewart, J. H. (1989). The knowledge structure of Mendelian genetics. *American Biology Teacher*, *51*(3), 143–149. https://doi.org/10.2307/4448880

Corbett, A., Kauffman, L., Maclaren, B., Wagner, A., & Jones, E. (2010). A cognitive tutor for genetics problem solving: Learning gains and student modeling. *Journal of Educational Computing Research*, *42*(2), 219–239. https://doi.org/10.2190/EC.42.2.e

DeCuir-Gunby, J. T., Marshall, P. L., & McCulloch, A. W. (2011). Developing and using a codebook for the analysis of interview data: An example from a professional development research project. *Field Methods*, *23*(2), 136–155. https://doi.org/10.1177/1525822X10388468

Freidenreich, H. B., Duncan, R. G., & Shea, N. (2011). Exploring middle school students' understanding of three conceptual models in genetics. *International Journal of Science Education*, *33*(17), 2323–2349. https://doi.org/10.1080/09500693.2010.536997

Hackling, M. W. (1994). Application of genetics knowledge to the solution of pedigree problems. *Research in Science Education*, *24*(1), 147–155. https://doi.org/10.1007/BF02356339

Hackling, M. W., & Lawrence, J. A. (1988). Expert and novice solutions of genetic pedigree problems. *Journal of Research in Science Teaching*, *25*(7), 531–546. https://doi.org/10.1002/tea.3660250703

Hickey, D. T., Wolfe, E. W., & Kindfield, A. C. H. (2000). Assessing learning in a technology-supported genetics environment: Evidential and systemic validity issues. *Educational Assessment*, *6*(3), 155–196. https://doi.org/10.1207/S15326977EA0603_1

Hofmann, H. (2008). Mosaic plots and their variants. In Chen, C.-H., Härdle, W., & Unwin, A. (Eds.), *Handbook of data visualization* (pp. 617–642). Berlin, Germany: Springer.

Knippels, M.-C. P. J., Waarlo, A. J., & Boersma, K. T. (2005). Design criteria for learning and teaching genetics. *Journal of Biological Education*, *39*(3), 108–112. https://doi.org/10.1080/00219266.2005.9655976

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., … & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115. https://doi.org/10.1016/j.edurev.2013.10.001

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lee, H.-S., Liu, O. L., Pallant, A., Roohr, K. C., Pryputniewicz, S., & Buck, Z. E. (2014). Assessment of uncertainty-infused scientific argumentation. *Journal of Research in Science Teaching*, *51*(5), 581–605. https://doi.org/10.1002/tea.21147

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences*, *15*(2), 153–191. https://doi.org/10.1207/s15327809jls1502_1

Ministry for School and Further Education of the State of North Rhine-Westphalia. (2014). *Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen. Biologie [Core curriculum for the upper secondary level Gymnasium/Gesamtschule in North Rhine-Westphalia. Biology]*. Düsseldorf, Germany. Retrieved January 21, 2022, from https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan/147/KLP_GOSt_Biologie.pdf

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, *19*, 1–13. https://doi.org/10.1177/1609406919899220

Sadler, T. D., & Fowler, S. R. (2006). A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education*, *90*(6), 986–1004. https://doi.org/10.1002/sce.20165

Sadler, T. D., & Zeidler, D. L. (2005). The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education*, *89*(1), 71–93. https://doi.org/10.1002/sce.20023

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction*, *23*(1), 23–55. https://doi.org/10.1207/s1532690xci2301_2

Schmiemann, P., Nehm, R. H., & Tornabene, R. E. (2017). Assessment of genetics understanding: Under what conditions do situational features have an impact on measures? *Science & Education*, *26*(10), 1161–1191. https://doi.org/10.1007/s11191-017-9925-z

Shea, N. A., Duncan, R. G., & Stephenson, C. (2015). A tri-part model for genetics literacy: Exploring undergraduate student reasoning about authentic genetics dilemmas. *Research in Science Education*, *45*(4), 485–507. https://doi.org/10.1007/s11165-014-9433-y

Smith, M. U. (1988). Successful and unsuccessful problem solving in classical genetic pedigrees. *Journal of Research in Science Teaching*, *25*(6), 411–433. https://doi.org/10.1002/tea.3660250602

Smith, M. U., & Good, R. (1984). Problem solving and classical genetics: Successful versus unsuccessful performance. *Journal of Research in Science Teaching*, *21*(9), 895–912. https://doi.org/10.1002/tea.3660210905

Songer, N. B., Kelcey, B., & Gotwals, A. W. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *Journal of Research in Science Teaching*, *46*(6), 610–631. https://doi.org/10.1002/tea.20313

Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany. (2005). *Bildungsstandards im Fach Biologie für den Mittleren* Schulabschluss [Educational standards in biology for the intermediate school leaving certificate]. München, Germany: Luchterhand. Retrieved January 21, 2022, from https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Biologie.pdf

Stewart, J. (1988). Potential learning outcomes from solving genetics problems: A typology of problems. *Science Education*, *72*(2), 237–254. https://doi.org/10.1002/sce.3730720211

Stewart, J., Cartier, J. L., & Passmore, C. M. (2005). Developing understanding through model-based inquiry. In National Research Council, *How students learn: History, mathematics, and science in the classroom* (pp. 515–565). Washington, DC: National Academies Press.

Surmann, J. (2017). *Teilautomatisierte Entwicklung und Analyse humangenetischer Stammbäume [Semi-automated development and analysis of human genetic pedigrees]* (Unpublished master's thesis). Germany: University of Duisburg-Essen.

Therneau, T. M., & Sinnwell, J. (2015). *Kinship2: Pedigree functions (R package version 1.6.4)*. Retrieved March 7, 2018, from https://CRAN.R-project.org/package=kinship2

Timm, J., Otto, B., Schramm, T., Striewe, M., Schmiemann, P., & Goedicke, M. (2020). Technical aspects of automated item generation for blended learning environments in biology. *I-Com*, *19*(1), 3–15. https://doi.org/10.1515/icom-2020-0001

Toulmin, S. E. (2003). *The uses of argument* (Updated ed.). Cambridge, England: Cambridge University Press. Retrieved from www.netLibrary.com/urlapi.asp?action=summary&v=1&bookid=120780

Tsui, C., & Treagust, D. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *International Journal of Science Education*, *32*(8), 1073–1098. https://doi.org/10.1080/09500690902951429

VERBI Software. (2018). *MAXQDA 2018: The art of data analysis. Berlin*.

von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, *45*(1), 101–131. https://doi.org/10.1002/tea.20213

Wools, K. (2018). *Strategien von Lernenden bei der Analyse von Familienstammbäumen [Strategies of learners in family tree analysis]* (Unpublished master's thesis). Germany: University of Duisburg-Essen.

Zhai, X., Shi, L., & Nehm, R. H. (2021). A meta-analysis of machine learning-based science assessments: Factors impacting machine-human score agreements. *Journal of Science Education and Technology*, *30*(3), 361–379. https://doi.org/10.1007/s10956-020-09875-z